Le raisonnement inductif sous toutes ses formes

Validité, limites, formes et aspects récents du raisonnement inductif

SOFIEN BEN AYED
Du 18 juin au 14 août 2018

Résumé

L'induction est un mode de raisonnement fondamental en mathématiques. Une fois démontré qu'il est possible d'étendre une relation d'un terme au terme qui lui succède, l'induction nous permet de généraliser cette relation à tous les termes de l'ensemble. C'est l'outil indispensable qui nous fait passer du fini à l'infini. Mais quels sont donc les fondements de ce raisonnement ? Pourquoi a-t-il déja fait débat ? À partir de quels axiomes sa validité peut-elle être prouvée ? Quand et sous quelles formes est-il applicable ?

Nous tentons de répondre à ces questions, et analysons les aspects récents de l'induction vue sous l'angle de l'énigmatique problème des chapeaux, dans les travaux de C. Hardin et A. Taylor. À partir de la connaissance d'un passé aussi proche que nous souhaitons, ils proposent une stratégie nous assurant une prédiction exacte presque sûrement du présent, et même de quelques instants futurs.

Mots-clés : *récurrence, induction, fondements des mathématiques, modèles de l'arithmétique, ω-incohérence, bon ordre, bonne fondation, axiome de l'infini, axiome du choix, problème des chapeaux, prédiction*

Table des matières

1 Le problème de l'induction **4**
D'où vient le problème ? . . . 4
Insoluble selon Hume . . . 4
Un pari risqué selon Russell . . . 4
Du point de vue des probabilités . . . 5
Poincaré, Goodman et la nouvelle énigme . . . 5
Est-ce vraiment un problème ? . . . 6

2 Le grand débat autour de l'induction au début du XX$^{\text{ème}}$ siècle **7**
Poincaré et l'induction . . . 7
La tentative de Russell . . . 7
Consolider l'arithmétique de Peano ? . . . 8

3 Équivalence entre l'induction et le principe du bon ordre **9**
Le bon ordre implique le principe d'induction . . . 9
Le principe d'induction implique le bon ordre . . . 9

4 Équivalence entre l'induction et le principe de bonne fondation **11**
Définitions et théorèmes . . . 11
La bonne fondation implique l'induction . . . 11
L'induction implique la bonne fondation . . . 12
Une généralisation de l'induction ? . . . 12

5 L'ω-incohérence, source d'interrogations autour du principe d'induction **13**
La conjecture de Goldbach peut-elle être indécidable ? . . . 13
Quelles conséquences aurait l'indécidabilité de la conjecture ? . . . 13
Quels sont les liens avec le principe d'induction ? . . . 14
Où en sommes-nous concernant la fameuse conjecture ? . . . 15

6 L'axiome d'induction : du 1$^{\text{er}}$ ou du 2$^{\text{nd}}$ ordre ? **16**
Énoncés de l'axiome d'induction . . . 16
La catégoricité . . . 16
La complétude . . . 17

7 Entre l'axiome du choix, le bon ordre et l'induction **18**
Énoncés de l'axiome du choix et du théorème du bon ordre . . . 18
L'axiome du choix implique le théorème du bon ordre . . . 18
Le théorème du bon ordre implique l'axiome du choix . . . 20
Quel lien avec l'induction ? . . . 20

8 L'induction comme conséquence de l'axiome de l'infini **21**
La théorie de Zermelo-Fraenkel . . . 21
Construction des entiers naturels . . . 22
L'induction en découle . . . 22

9 Le problème des chapeaux à l'infini et son rapport à l'induction **23**
L'énoncé géneral du problème . . . 23
La stratégie Hardin-Taylor . . . 23
Dans le cas dénombrable . . . 24
Dans le cas indénombrable . . . 24
Prédire le futur ? . . . 25
De quelle sorte d'induction s'agit-il ? . . . 26
Le problème de l'induction est-il résolu ? . . . 26
Alexander George . . . 26
Alexander Paseau . . . 27

10 Réflexions, problèmes ouverts et perspectives **29**

Annexe A Équivalence entre l'induction simple et forte **34**
Les énoncés des principes . . . 34
L'induction forte implique l'induction simple . . . 34
L'induction simple implique l'induction forte . . . 34

1 Le problème de l'induction

Résumé

En sciences physiques et naturelles, la connaissance est acquise par un raisonnement inductif : il consiste à remonter de cas particuliers à la loi qui les régit. Or, la rigueur de ce raisonnement n'a jamais pu être justifiée : il s'agit du problème de l'induction. C'est le philosophe Hume qui l'a énoncé pour la première fois, et depuis, plusieurs réponses ont éte apportées pour tenter de justifier l'induction.

D'où vient le problème ?

Est-ce que le raisonnement inductif conduit à la connaissance ? C'est la question principale posée par le problème de l'induction. Il est possible de distinguer deux questions plus précises [Coz09] :

- Peut-on généraliser des propriétés sur la base d'observations répétées ? Si tous les cygnes qu'on a vu jusqu'à présent sont blancs, est-ce que cela veut dire que tous les cygnes sont blancs ?
- Est-il juste de supposer qu'une séquence d'évènements à l'avenir se produira comme elle l'a toujours fait dans le passé ? Est-ce qu'une boule qui en cogne une autre va toujours réagir de la même manière ? C'est ce que Hume appelle le principe de l'uniformitarisme.

Etant donné que le raisonnement inductif est à la base du savoir scientifique (en particulier les sciences physiques et naturelles), remettre en cause sa validité, c'est remettre en cause quasiment toutes les certitudes auxquelles a abouti la science.

Insoluble selon Hume

David Hume a remis en question l'induction dans ***Enquête sur l'entendement humain* en 1748**. Selon lui, l'induction n'a pas un fondement logique mais psychologique. Nous avons l'habitude que le soleil se lève tous les jours donc aucune raison de penser qu'il ne se lèvera pas demain. Oui mais il est impossible de savoir si la Nature est vraiment uniforme. D'un point de vue seulement logique, si tout s'est toujours produit en suivant certaines règles, ça n'implique pas que les règles ne peuvent pas changer demain. Le futur ne ressemble au passé que par habitude.

Un pari risqué selon Russell

Bertrand Russell compare le raisonnement inductif au raisonnement que tient le poulet lorsqu'il associe la main du fermier au grain qui le nourrit, jusqu'au jour où cette même main lui tordra le cou. L'induction est vue comme un pari risqué.

Et quand bien même la Nature serait uniforme, quel que soit le nombre de fois où le même phénomène est expérimenté, il est impossible d'épuiser tous les cas. En d'autres termes, il n'existe pas de test exhaustif sur une infinité de cas.

Pourtant, les théories qui ont été établies sur la base d'un raisonnement inductif n'ont pas été remises en question. Du moins, ce n'est jamais ce raisonnement qui est mis en défaut lorsqu'une théorie est critiqué. Nous pourrions donc affirmer que si l'induction ne nous a jamais fait défaut jusqu'aujourd'hui, alors il en sera toujours de même. Comme le soleil qui se lèvera demain, l'induction fonctionnera encore demain. Mais il s'agit là d'une tentative de justification de l'induction par l'induction elle-même. C'est un raisonnement circulaire qui ne peut fonctionner [Hen18].

Du point de vue des probabilités

Il n'est pas question d'abandonner l'induction puisque nous y faisons appel systématiquement mais il est possible de la limiter. Si un grand nombre n d'étoiles sont observées et qu'aucune ne possède la propriété P, on ne dira pas qu'aucune étoile ne la possède mais seulement que la probabilité que la prochaine étoile étudiée ait la propriété P est inférieure à $\frac{1}{n+2}$ selon la loi de succession de Laplace. C'est une induction dite bornée. Si nous nous limitons à ce genre d'induction alors nous n'aboutissons plus à aucune certitude [Hen18].

Poincaré, Goodman et la nouvelle énigme

D'autres ont reformulé le problème de l'induction comme **Nelson Goodman en 1946** avec sa nouvelle énigme de l'induction. Il invente l'adjectif « *vleu* » qui signifie vert jusqu'à une certaine date t et bleu ensuite. De cette manière, si toutes les émeraudes observées sont vertes, il est envisageable d'en conclure par induction soit que toutes les émeraudes sont vertes, soit que toutes les émeraudes sont vleues [LY99]. Pourquoi l'une de ces deux affirmations aurait plus de chance d'être valide ? En réalité, cette nouvelle énigme de l'induction apparaissait déjà 50 ans auparavant dans **les écrits de Poincaré portant sur l'interpolation des fonctions et l'induction** [Poi02].

« *Je veux déterminer une loi expérimentale ; cette loi, quand je la connaîtrai, pourra être représentée par une courbe ; je fais un certain nombre d'observations isolées ; chacune d'elles sera représentée par un point. Quand j'ai obtenu ces différents points, je fais passer une courbe entre ces points en m'efforçant de m'en écarter le moins possible et, cependant, de conserver à ma courbe une forme régulière, sans points anguleux, sans inflexions trop accentuées, sans variation brusque du rayon de courbure. Cette courbe me représentera la loi probable, et j'admets, non seulement qu'elle me fait connaître les valeurs de la fonction intermédiaires entre celles qui ont été observées, mais encore qu'elle me fait connaître les valeurs observées elles-mêmes plus exactement que l'observation directe.*

Pourquoi donc est-ce que je cherche à tracer une courbe sans sinuosités ? C'est parce que je considère a priori une loi représentée par une fonction continue (ou par une fonction dont les dérivées d'ordre élevé sont petites), comme plus probable qu'une loi ne satisfaisant pas à ces conditions. Sans cette croyance, le problème dont nous parlons n'aurait aucun sens ; l'interpolation serait impossible ; on ne pourrait déduire une loi d'un nombre fini d'observations ; la science n'existerait pas. »

Comme dans le cas de l'énigme de Goodman, ce qui est en question est bien l'objet du choix, et non sa plus ou moins grande probabilité de voir ce choix couronné de succès. Et selon Poincaré, il faut aller au plus simple. Cette notion de simplicité peut être rapprochée du rasoir d'Ockham formulé au XIVème siècle qui peut se peut se traduire comme : « *les hypothèses les plus simples sont les plus vraisemblables* ».

Est-ce vraiment un problème ?

Si le problème de l'induction est insurmontable, pourquoi ne pas tenter de vivre avec ? Selon le philosophe des sciences **Karl Popper**, le problème de l'induction ne pose pas autant de problèmes qu'il n'y parait à la science. **En effet, la science ne cherche pas à justifier des théories par induction mais repose sur des conjectures et des critiques. La science cherche des théories qui sont possiblement fausses mais dont toutes les tentatives pour les réfuter ont échoué jusqu'à présent.** [Hen18] C'est pourquoi la théorie des forces newtonienne a fait son chemin avant d'être mise en défaut par Einstein.

Enfin, en mathématiques, le problème de l'induction tel qu'il a été formulé par Hume ne se pose pas de la même manière puisqu'il s'agit avant tout de ce que Hume appelle des relations d'idées et non des faits comme en sciences physiques et naturelles. Notamment, le problème de l'uniformité de la Nature ne se pose pas.

2 Le grand débat autour de l'induction au début du XX$^{\text{ème}}$ siècle

Résumé

Dans le cadre du grand débat autour des fondements des mathématiques qui a eu lieu au début du XX$^{\text{ème}}$ siècle, le principe d'induction est remis en question. Comment légitimer ce principe ? C'est la question que se sont posés Poincaré, les logicistes menés par Russell ainsi que les axiomatistes Peano et Hilbert.

Poincaré et l'induction

Un raisonnement inductif peut-il être rigoureux ? C'est la question que Poincaré se pose dans le premier chapitre de *La Science et l'hypothèse*, intitulé « *Sur la nature du raisonnement mathématique* ». Si cette question revêt une telle importance, c'est parce que selon lui, le seul raisonnement déductif est impuissant. La science a besoin de l'induction mais elle doit aussi être rigoureuse. Tous ne sont pas de cet avis concernant le raisonnement déductif. Pour Frege, les conclusions sont effectivement contenues dans les définitions mais cela ne rend pas pour autant ce raisonnement stérile car « *elles le sont comme une plante l'est dans une graine* ».

Poincaré prend l'exemple du raisonnement par récurrence, non déductif et pourtant indispensable pour démontrer bon nombre de théorèmes généraux de l'arithmétique. Le principe d'induction qui fonde le raisonnement par récurrence ne peut ni être donné par l'expérience, ni se réduire à une convention comme les axiomes de la géométrie, ni même être rapprochée à l'induction en sciences physiques et naturelles qui est systématiquement incertaine. Il est donc difficile d'extraire la substance de ce raisonnement alors même qu'il parait évident. Poincaré parle de « *l'intuition du nombre pur* », notre capacité à concevoir sans mal la répétition indéfinie.

La tentative de Russell

Les logicistes Frege et Russell ont eux tenté de donner un fondement purement logique à l'arithmétique et donc au principe d'induction qui était jusqu'alors le 5ème axiome dans la théorie des nombres de Peano. Rappelons d'ailleurs ces 5 axiomes de Peano [Col09] :

Axiomes 1 (Peano). *L'axiomatique des nombres selon Peano :*

1. 0 *est un nombre.*
   $0 \in \mathbb{N}$
2. *Si* n *est un nombre alors le successeur de* n *est un nombre.*
   $n \in \mathbb{N} \Rightarrow S(n) \in \mathbb{N}$
3. 0 *n'est le successeur d'aucun nombre.*
   $\forall n \in \mathbb{N}, S(n) \neq 0$
4. *Deux nombres qui ont le même successeur sont égaux.*
   $\forall n, m \in \mathbb{N}, S(n) = S(m) \Rightarrow n = m$

5. *Si une propriété est vérifiée pour 0 et si, pour tout entier naturel* n *qui la vérifie,* $S(n)$ *la vérifie également, alors la propriété est vraie pour tous les entiers naturels.*

Pour y parvenir, Russell définit 5 nouvelles notions :

Propriété héréditaire - si elle appartient à un nombre, elle appartient aussi à son successeur.

Classe héréditaire - si n appartient à cette classe, n+1 aussi.

Propriété inductive - propriété héréditaire et qui appartient à 0.

Classe inductive - classe héréditaire et qui contient 0.

Postérité d'un nombre - tous les objets qui appartiennent à toute classe héréditaire contenant ce nombre.

Les nombres naturels sont alors définis comme la postérité de 0. Cette définition permet à Russell de ne garder que 3 des 5 propriétés primitives de Peano : la 1ère qui affirme que 0 est un nombre et le principe d'induction ne sont que des conséquences de la définition. En effet, il est inutile de mentionner que 0 est un nombre puisque les nombres sont définis à partir de ce 0 et l'incluent (la postéritè de 0). Pour ce qui est de l'induction, elle est contenue dans la notion de classe héréditaire à partir de laquelle est définie la postérité d'un nombre.

En 1905, Poincaré critique les travaux de Russell. Son objection majeure concerne l'usage du mot « *tous* ». Quand le nombre d'objets est infini, le « *tous* » perd son sens. La notion de postérité dépend de *tous les objets* or nous ne pouvons pas définir tous ces objets sans faire appel à la notion de postérité elle-même. C'est un cercle vicieux que reconnaît Russell. Après de nombreuses autres tentatives infructueuses, Russell avoua son échec en 1925. La logique n'a pas suffi à fonder l'arithmétique.

Consolider l'arithmétique de Peano ?

Poincaré pensait tout de même comme Russell que le système de Peano était insuffisant. En effet, tous deux lui reprochaient de ne pas avoir prouvé la non-contradiction de son système axiomatique, preuve que Poincaré considère « *aussi nécessaire qu'impossible* ». Hilbert se rapproche de Poincaré en ce sens mais lui pense que la preuve est possible et tente de l'effectuer. D'autres s'y sont aussi essayés par la suite tels que Ackermann, Herbrand ou encore von Newmann.

Toutes ces tentatives ont conduit à la construction de systèmes axiomatiques de plus en plus satisfaisants mais en 1931, Gödel y met un terme. Il démontre qu'il est impossible de prouver la non-contradiction de l'arithmétique à l'intérieur de ce système. C'est son second théorème d'incomplétude.

Théorème 1 (Incomplétude). *[Göd92] Si* T *est une théorie du premier ordre cohérente, récursivement axiomatisable et contenant l'arithmétique de Peano, alors la cohérence de la théorie T n'est pas une conséquence des axiomes de T.*

La preuve de non-contradiction n'est donc possible que si on ne se limite pas aux outils strictement arithmétiques [Bon04].

3 Équivalence entre l'induction et le principe du bon ordre

Résumé

A la recherche d'une démonstration de l'induction, le résultat trouvé le plus connu est une preuve faisant appel au principe du bon ordre. Il y a même équivalence entre les deux principes et en voici la démonstration.

Le bon ordre implique le principe d'induction

Dans l'arithmétique de Peano, les entiers ont été définis comme étant soit 0 soit $n+1$, c'est à dire le successeur d'un autre entier. Cette affirmation sera utile dans les deux preuves que nous allons mener.

Théorème 2. *Si P est une proposition sur les entiers naturels telle que :*

- $P(0)$ *est vrai.*
- $\forall m, (P(m) \Rightarrow P(m+1))$.

Alors P *est vraie pour tous les entiers naturels.*

Démonstration. Supposons que l'ensemble des entiers naturels est bien ordonné selon la relation d'ordre, $\forall n \in \mathbb{N}, n < n+1$.

Soit S l'ensemble de tous les entiers naturels n pour lesquels $P(n)$ est faux. Supposons que S contient au moins un élément.

Alors d'après le principe de bon ordre, S a un plus petit élément n. De plus, puisque $P(0)$ est vrai, n ne peut pas être 0. Aussi, puisque chaque entier naturel est soit 0 soit $m+1$ avec m un entier naturel, $n = m+1$. Ainsi, $m < n$.

Or, n est le plus petit élément de S donc m n'est pas dans S et $P(m)$ est donc vrai. Alors, par hypothèse, $P(m+1)$ qui correspond à $P(n)$ est aussi vrai car $n = m+1$. C'est une contradiction car n est dans S, donc $P(n)$ est faux.

Donc S est vide, c'est-à-dire que P est vraie pour tous les entiers naturels.

□

Le principe d'induction implique le bon ordre

Nous voulons montrer que tout sous-ensemble non-vide admet un plus petit élément. Dans la démonstration, nous aurons besoin d'employer le principe d'induction forte, un principe qui est équivalent à l'induction simple telle que nous l'avons définie jusqu'à présent. Une preuve en est dressée en annexe A. L'induction forte s'énonce comme suit :

Théorème 3. *Soit* P *une propriété définie sur* $\mathbb{N}$. *Si :*

- $P(0)$ *est vrai.*
- $\forall n \in \mathbb{N}, [\forall k \leqslant n, P(k)] \Rightarrow P(n+1)$

Alors $\forall n, P(n)$ *est vraie.*

Démonstration. Supposons que le principe d'induction énoncé précédemment est vrai.

Soit $S \subset \mathbb{N}$ un ensemble non-vide. Supposons que S n'admet pas de plus petit élément.

D'abord, si $0 \in S$, sachant qu'il n'y a pas d'entier naturel inférieur à 0, alors S admet 0 comme plus petit élément ce qui contredit notre hypothèse.

Maintenant, intéréssons-nous au cas où $0 \notin S$. On prouve par induction forte que S est vide, soit que $\forall n \in \mathbb{N}, n \notin S$. On note cette propriété P.

Comme $0 \notin S$, $P(0)$ est vrai.

Supposons que $P(k)$ est vraie $\forall k \leqslant n$. Cela signifie que $\forall k \leqslant n, k \notin S$. Supposons à nouveau deux cas.

Si $n+1 \in S$ alors S admet $n+1$ comme plus petit élément ce qui contredit notre hypothèse.

Et si $n+1 \notin S$, alors $P(n+1)$ est vrai. Par induction forte, nous pouvons donc conclure que $\forall n, P(n)$ est vrai. En d'autres termes, $\forall n \in \mathbb{N}, n \notin S$. Donc S est vide ce qui est contradictoire.

Dans tous les cas, nous aboutissons à une contradiction. Donc tout sous-ensemble S non-vide admet un plus petit élément.

$\square$

Cette relation sur les entiers peut être étendue à tout ensemble bien ordonné. Il y a donc équivalence entre le principe d'induction et le principe de bon ordre. **Mais ce qui nous intéresse davantage, c'est de savoir s'il est possible de déduire le principe d'induction d'un ensemble d'axiomes plus faibles que le bon ordre.**

4 Équivalence entre l'induction et le principe de bonne fondation

Résumé

Nous avons démontré précédemment qu'il ést possible de raisonner par induction sur tout ensemble bien ordonné. Un autre principe d'induction légèrement différent permet de travailler avec une condition affaiblie par rapport à la condition de bon ordre : c'est la condition de bonne fondation. Montrons que ce nouveau principe d'induction est équivalent à la bonne fondation.

Définitions et théorèmes

Un élément minimal d'un ensemble muni d'un ordre strict est défini comme suit [Liv] :

> **Définition 1.** *Soit* A *un ensemble muni d'un ordre strict* $\prec$*, c'est à dire une relation d'ordre irréflexive et transitive. L'élément* $x \in A$ *est dit minimal si et seulement s'il n'existe aucun autre élément de cet ensemble qui lui soit inférieur.*

Les ordres bien fondés sont caractérisés comme suit :

> **Définition 2.** *Soit* A *un ensemble ordonné selon un ordre* $<$*. L'ordre est bien fondé si et seulement si toute partie non vide* B *de* A *admet un élément minimal pour* $<$*.* A *est alors dit bien fondé.*

Le principe d'induction bien fondé peut alors s'écrire comme suit :

> **Principe 1.** *Soient* A *un ensemble muni d'un ordre strict bien fondé et* P *une propriété qui concerne les éléments de cet ensemble. Si :*
>
> 1. *pour tous les éléments minimaux* m*,* $P(m)$ *est vrai,*
> 2. *si* $\forall k \prec x$*,* $P(k)$ *est vrai alors* $P(x)$ *est vrai,*
>
> *alors* $\forall x \in A, P(x)$ *est vrai.*

Nous voulons démontrer l'équivalence entre la validité de ce principe d'induction bien fondé et la bonne fondation d'un ensemble.

La bonne fondation implique l'induction

Démontrons que la bonne fondation d'un ensemble implique la validité du principe d'induction sur ce même ensemble [Mye13].

Démonstration. Soient A un ensemble muni d'un ordre bien fondé. Supposons que le principe d'induction n'est pas vérifié.

Il existe une propriété P vérifiée pour tous les éléments minimaux de A et telle que si $\forall k \prec x$, $P(k)$ est vrai, alors $P(x)$ est vrai, mais telle que l'affirmation $\forall x \in A, P(x)$ est fausse. Il existe donc au moins un élément $a \in A$ tel que $P(a)$ est faux.

Définissons l'ensemble $B = \{x \in A \mid P(x) est faux\}$. Cet ensemble est non-vide. Puisque A est bien fondé, B admet un élément minimal que nous appelons x_0. Donc $\forall y < x_0, y \notin B$, d'où $\forall y < x_0, P(y)$ est vrai. Ceci implique que $P(x_0)$ est vrai. Nous avons abouti à une contradiction, car par définition de B, $P(x_0)$ est faux.

Le principe d'induction est donc vérifié sur un ensemble muni d'un ordre bien fondé.

□

L'induction implique la bonne fondation

Démontrons que la validité du principe d'induction sur un ensemble implique la bonne fondation de cet ensemble.

Démonstration. Nous supposons le principe d'induction vérifié sur un ensemble A. Supposons que cet ensemble n'est pas bien fondé.

Il existe un ensemble $B \subset A$ non-vide tel que B ne contienne pas d'élément minimal.

Posons la propriété $P(x) =$ "$x \notin B$". B ne peut pas contenir un élément minimal de A car B n'a pas d'élément minimal et $B \subset A$. Donc pour tout élément minimal $a \in A$, $a \notin B$, donc $P(a)$ est vrai.

Supposons maintenant que $\forall k \prec x$, $P(k)$ est vrai. Ceci signifie que $\forall k \prec x, k \notin B$. Donc si $x \in B$ alors x est un élément minimal de B ce qui contredit notre hypothèse. Donc $x \notin B$ d'où $P(x)$ est vrai. Par induction, nous concluons que $\forall x \in A, P(x)$ est vrai. En d'autres termes, $\forall x \in A, x \notin B$. L'ensemble B est donc vide.

Nous avons abouti à une contradiction. L'ensemble sur lequel le principe d'induction est vérifié, est donc bien fondé.

□

Une généralisation de l'induction ?

L'équivalence entre le bon ordre et l'induction nous permet de raisonner par induction sur des ensembles bien ordonnés. Nous sommes ici passé à un cas plus général qui nous permet de raisonner par induction sur des ensembles bien fondés. En quoi s'agit-il d'une généralisation ? Un bon ordre est toujours bien fondé mais la réciproque n'est pas vraie. **En effet, par définition, un bon ordre est un ordre à la fois bien fondé et total, c'est à dire si deux éléments sont toujours comparables.**

Cette généralisation peut s'avérer très utile par exemple dans le cas d'un ensemble de structures de données construites de manière récursive. Elle a même un nom dans ce cas-ci : induction structurelle. Mais nul besoin d'aller bien loin pour trouver un emploi utile de l'induction bien fondée. Munissons à l'ensemble des entiers naturels la relation d'ordre R définie par aRb si et seulement si a divise b et $a \neq 1$. R est bien fondé mais pas total. Le premier volet du théorème fondamental de l'arithmétique concerne l'existence d'une décomposition en facteurs premiers pour tout entier strictement positif (le deuxième concernant son unicité). Cette existence peut être démontrée par induction bien fondée.

5 L'ω-incohérence, source d'interrogations autour du principe d'induction

Résumé

Si pour chaque entier, une théorie peut prouver qu'une propriété est vraie, cela n'implique pas que cette théorie peut prouver que la même propriété est vérifiée pour tout entier, et c'est pour cette raison que nous avons besoin du principe d'induction. A travers l'exemple de la conjecture de Goldbach et les conséquences de sa potentielle indécidabilité, nous tentons de mieux comprendre la raison de cette non-implication pour finalement faire le lien avec le principe d'induction et les difficultés qu'il pose.

La conjecture de Goldbach peut-elle être indécidable ?

Les affirmations mathématiques peuvent se distinguer par leur forme logique. Par exemple, le premier niveau appelé Δ_0 concerne les affirmations que l'on peut vérifier. Ainsi, l'affirmation "36 est la somme de deux nombres premiers" est de ce type. Il est aisé de la vérifier en effectuant toutes les sommes possibles de nombres premiers jusque 36. Dans l'arithmétique de Peano, toutes les affirmations de Δ_0 qui sont vraies sont démontrables. Si elles sont fausses, il est également possible de le démontrer. Autrement dit, aucune affirmation de Δ_0 n'est indécidable.

La conjecture de Goldbach se formule comme suit :

> **Conjecture 1.** *Tout nombre entier pair strictement supérieur à 2 peut s'écrire comme la somme de deux nombres premiers.*

C'est une conjecture qui a été formulée par Christian Goldbach en 1742 et dont aucune démonstration n'a été trouvée jusqu'à présent.

Il s'agit d'une affirmation dite Π_1, car elle est de la forme "$\forall n P(n)$" où $P(n)$ est une affirmation de Δ_0. Or, certaines affirmations Π_1 sont indécidables. Il n'est donc pas impossible que la conjecture de Goldbach soit indécidable et l'arithmétique de Peano serait alors incapable de se prononcer sur cette conjecture : ni prouvable, ni réfutable [Del14].

Quelles conséquences aurait l'indécidabilité de la conjecture ?

Il est possible qu'un jour soit prouvé, dans une théorie plus forte comme la théorie des ensembles, l'indécidabilité de la conjecture de Goldbach dans l'arithmétique de Peano (AP). Cela voudrait également dire que la conjecture de Goldbach est vraie. Il est ici nécessaire de définir cette notion de véracité. Une formule est dite vraie (ou satisfaisable) s'il existe une interprétation, plus précisément un modèle dans lequel cette formule est vérifiée. Elle est dite valide si elle est vérifiée dans tous les modèles.

Pourquoi l'indécidabilié de la conjecture de Goldbach implique sa véracité ? En fait, si elle était fausse alors cela signifierait qu'il existe un nombre pair qui ne peut pas s'écrire comme la somme de deux nombres premiers. Mais si ce contre-exemple existait alors nous pourrions le trouver mécaniquement et réfuter la conjecture ce qui en contradiction avec son indécidabilité. De ce fait, si son indécidabilité est prouvée alors elle doit être vraie [Kac09].

En définitive, il se pourrait que la conjecture de Goldbach soit à la fois indécidable et vraie. Exhibons deux modèles : l'un qui prend comme nouvel axiome la véracité de la conjecture $\mathsf{AP} + (\mathtt{Goldbach})$ et un autre qui prend sa non-véracité $\mathsf{AP} + \neg(\mathtt{Goldbach})$. En fait, la conjecture de Goldbach est vraie uniquement pour les nombres dits standards, c'est à dire tous ceux que nous pouvons atteindre par additions successives. Ici, $\mathsf{AP} + \neg(\mathtt{Goldbach})$ est donc nécessairement un modèle non-standard de l'arithmétique dans lequel il existe un nombre non-standard qui ne peut pas s'écrire comme la somme de deux nombres premiers non-standards [Kem58]. C'est Albert Skolem en collaboration avec Jacques Herbrand qui a introduit cette notion de modèle non-standard de l'arithmétique [Sko34]. Il s'agit d'un système mathématique qui vérifie toutes les propriétés élémentaires des nombres entiers, mais qui diffère du modèle standard. Pour construire un tel modèle, il est par exemple possible de se contenter d'ajouter aux axiomes de Peano une nouvelle constante c définie de la manière suivante :

> **Définition 3.** c *est définie par le schéma d'axiomes* $\forall n \in \mathbb{N}, n < c$.

Il existe bien sûr d'autres modèles non-standards dont la construction est plus complexe, notamment celle des hypernaturels [Dav09].

Quels sont les liens avec le principe d'induction ?

Dans le cadre de la preuve de son premier théorème d'incomplétude, Kurt Gödel a introduit la notion d'ω-cohérence.

> **Définition 4.** *[Ana02] Une théorie* T ***arithmétique*** *est dite* ω*-incohérente si il existe une propriété portant sur les entiers naturels telle que* $\forall n$, $T \vdash P(n)$ *(c'est à dire pour tous les entiers naturels standards) et telle qu'il existe une démonstration de l'existence d'un entier (nécessairement non-standard) pour lequel* $P(n)$ *est faux.*

C'est cette notion d'ω-incohérence qui rend le principe d'induction indispensable comme nous allons le montrer.

Posons les conditions d'application du raisonnement par récurrence. Soit P une propriété portant sur les entiers naturels tel que $P(0)$ est vrai et que $\forall n[P(n) \Rightarrow P(n+1)]$. Sans le principe d'induction admis comme axiome dans l'arithmétique de Peano, nous pouvons seulement conclure que $\forall n$ il existe une démonstration de $P(n)$. En effet, la démonstration pour chaque n consiste en une simple suite de n implications successives de $P(0)$ à $P(n)$. En revanche, il nous est impossible d'affirmer qu'il existe une démonstration de $\forall n, P(n)$. C'est la présence d'énoncés vrais de la forme Π_1 mais indécidables dans l'arithmétique de Peano comme pourrait peut-être l'être la conjecture de Goldbach qui en est la cause. La théorie est ω-incohérente. Ainsi, si nous avons besoin du principe d'induction, c'est parce que démontrer une propriété pour chaque n dans ce genre de théorie ne suffit pas à montrer que $\forall n, P(n)$.

Nous pouvons aller plus loin et montrer que dans certains cas bien précis, le principe d'induction peut être appliqué à tort à cause de l'existence des modèles non-standards mentionnés précédemment.

Prenon un exemple. Soit la propriété P, "n est fini".

> **Définition 5.** n *est fini si et seulement si il existe* m *fini tel que* $n = S(m)$ *avec* $S(m)$ *le successeur de* m *ou si* $n = 0$.

$P(0)$ est vrai puisque 0 est fini. P est aussi héréditaire car si n est fini alors $S(n)$ aussi. Par récurrence, nous voudrions donc conclure que $\forall n, n$ est fini. Or, cette conclusion est fausse car il existe des entiers non-standards supérieurs à tous les entiers standards, ils sont dits infiniment grands. Aucun entier non-standard ne peut être atteint à partir d'un entier standard par applicaiton du successeur. La frontière est extrêmement floue entre les standards et les non-standards [Sal88]. C'est ce qui nous conduit à être très prudent vis-à-vis de l'application du principe d'induction. Ce résultat fait écho à ce qu'a écrit Henri Poincaré au début du XX$^{\text{ème}}$ siècle [Poi02] :

« *Le principe d'induction ne signifie pas que tout nombre entier peut être obtenu par additions successives ; il signifie que, pour tous les nombres que l'on peut obtenir par additions successives, on peut démontrer une propriété quelconque par voie de récurrence.* »

Par prudence, on peut écrire le principe de récurrence restreint suivant (dit aussi pragmatique) [Del02] :

> **Théorème 4.** *Si* $P(0)$ *est vrai, et si pour tout* n *standard,* $P(n) \Rightarrow P(n+1)$*, alors pour tout* n *standard,* $P(n)$ *est vérifié.*

En conclusion, pour que $\forall n \in \mathbb{N}, T \vdash P(n)$ implique $T \vdash \forall n\, P(n)$, T doit être ω-cohérente.

Où en sommes-nous concernant la fameuse conjecture ?

Elle a été vérifiée pour tous les entiers pairs jusqu'à 4.10^{18} [eSHP14]. Et du point de vue théorique, en 1995, Olivier Ramaré a prouvé que tous les nombres pairs strictement supérieurs à 2 peuvent s'écrire comme la somme d'au plus 6 nombres premiers [Ram95]. Peut-être que la conjecture de Goldbach sera un jour prouvée indécidable ? C'est ce que certains ont pensé du dernier théorème de Fermat avant qu'il soit bel et bien démontré...

6 L'axiome d'induction : du 1^{er} ou du 2^{nd} ordre ?

Résumé

Le $5^{ème}$ axiome de Peano qui nous occupe depuis le début peut être formulé de deux manières différentes : soit sous la forme d'un schéma d'axiomes du 1^{er} ordre soit d'un unique axiome du 2^{nd} ordre. En apparence, cela ne fait guère de différences. Pourtant, nous allons voir que les propriétes des nombres de Peano sont radicalement différentes suivant la formulation employée.

Énoncés de l'axiome d'induction

L'arithmétique de Peano peut s'écrire sous deux formes : l'une est une formulation du 1^{er} ordre et l'autre du 2^{nd} ordre. C'est l'axiome d'induction qui permet ces deux formulations. Les axiomes de Peano ont été rappelés au chapitre 2. Venons-en donc directement au $5^{ème}$ axiome.

Il est possible d'exprimer l'induction par un schéma d'axiomes du 1^{er} ordre, c'est à dire par une infinité d'axiomes de la forme suivante, avec P un prédicat du 1^{er} ordre :

$$(P(0) \wedge \forall n, (P(n) \Rightarrow P(n+1))) \Rightarrow \forall n, P(n)$$

Il est également possible de l'exprimer par un unique prédicat du 2^{nd} ordre. En effet, la logique du 2^{nd} ordre autorise la quantification des prédicats. Par extension de la formulation du 1^{er} ordre, on écrit :

$$\forall P, ((P(0) \wedge \forall n, (P(n) \Rightarrow P(n+1))) \Rightarrow \forall n, P(n))$$

En apparence, il ne semble pas y avoir de différences fondamentales entre les deux. Pourtant, formuler l'induction d'une manière ou de l'autre a d'importantes conséquences sur les propriétés de la théorie axiomatique de que l'on bâtit.

La catégoricité

Le premier point de divergence que nous abordons entre les deux formulations concerne leur catégoricité.

> **Définition 6.** *[MP89] Une théorie T est catégorique si deux modèles quelconques M_i et M_j de T sont toujours isomorphes. En termes ensemblistes, un isomorphisme est une bijection.*

Plus simplement, une théorie catégorique n'admet donc qu'un seul modèle.

La formulation du 1^{er} ordre de l'arithmétique de Peano n'est pas catégorique. Il existe en effet des modèles qui ne sont pas isomorphes au modèle standard des entiers naturels. Nous avons vu qu'il est possible d'exhiber des modèles non-standards de l'arithmétique dans lesquels certains nombres sont plus grands que tous les nombres dits standards. Par exemple, avec une nouvelle constante c et le schéma d'axiomes $n < c$ pour tous les entiers n. Chaque sous-ensemble fini des axiomes est bien vérifié dans le modèle standard de l'arithmétique. Employons le théorème de compacité mais rappelons avant la définition de la satisfiabilité :

Définition 7. *Une théorie* T *est satisfiable si elle admet au moins un modèle. Elle est finiment satisfiable si toute partie finie* $T_f \subseteq T$ *est satisfiable.*

Le théorème de compacité s'énonce comme suit :

Théorème 5. *[Cag15] Une théorie est satisfiable si et seulement si elle est finiment satisfiable.*

D'après ce théorème, il y a bien un modèle qui satisfait l'ensemble des axiomes de Peano enrichi par le schéma d'axiomes $n < c$. Nous venons donc bien d'exhiber un modèle non-standard qui comprend un nombre non-standard qui interprète c.

Nous savons de toute façon, par le théorème de Löwenheim-Skolem, qu'aucune théorie du 1er ordre qui admet un modèle infini n'est catégorique [Lad69].

Au contraire, la formulation du 2nd ordre est catégorique. Dedekind en a dressé la preuve en 1888 [Ded65]. Tous les modèles de l'arithmétique formulée au 2nd ordre sont isomorphes. Comme dans toutes les théories catégoriques, **les notions de véracité et de validité sont confondues : il n'existe pas de formule à la fois vraie dans un modèle et fausse dans un autre.** Ainsi, contrairement à l'axiomatique du 1er ordre pour laquelle il existe des modèles non-standards des entiers, l'axiomatique du 2nd ordre décrit uniquement les entiers naturels tels que nous les connaissons, c'est à dire $\mathbb{N}$.

La complétude

De prime abord, la catégoricité de la formulation du 2nd ordre semble pratique de par l'unicité, à l'isomorphisme près de son modèle. Néanmoins, cette formulation présente un mal bien plus profond que la non-catégoricité, qui réside dans son système de déduction.

Le théorème de complétude de Gödel ne s'applique qu'à la logique du 1er ordre et donc pas à notre théorie axiomatique du 2nd ordre. Rappelons ce thèorème démontré en 1920 par Kurt Gödel :

Théorème 6. *[Göd29] Si* T *est une théorie de la logique du premier ordre, alors* T *est complète. En d'autres termes, soit une formule* φ *de la logique du premier ordre. Si* φ *est conséquence sémantique de* T *alors* φ *est conséquence syntaxique de* T*. En d'autres termes,* ***tout énoncé valide, c'est à dire vrai dans tous les modèles****, peut être démontré.*

Tout énoncé valide est prouvable. En ce sens, l'arithmétique formulée au 1er ordre est donc bien complète. Mais il n'en est rien pour la formulation du 2nd ordre.

En effet, le modèle standard de l'arithmétique de Peano formulée au 2nd ordre est infini et unique du fait de sa catégoricité. Or, d'après le théorème de Löwenheim-Skolem, une telle théorie ne peut pas être complète [Rag16].

C'est cette non-complétude de l'arithmétique de Peano avec l'axiome d'induction formulé comme un prédicat du 2nd ordre qui nous fait largement préferer le schéma d'axiomes de 1er ordre même si la théorie ne caractérise pas uniquement les entiers usuels.

7 Entre l'axiome du choix, le bon ordre et l'induction

Résumé

Toujours dans l'esprit de vouloir affaiblir les conditions nécessaires à l'application de l'induction, pourrions-nous raisonner par induction sur n'importe quel ensemble ? Oui, à condition d'utiliser l'axiome du choix, longtemps controversé mais aujourd'hui largement accepté comme un axiome à part entière de la théorie des ensembles. En effet, il est équivalent au théorème du bon ordre qui affirme que tout ensemble peut être bien ordonné. C'est d'ailleurs sur ce principe que repose la solution au problème des chapeaux de Hardin et Taylor, à laquelle nous nous intéresserons au chapitre suivant.

Énoncés de l'axiome du choix et du théorème du bon ordre

L'axiome du choix est un axiome de la théorie des ensembles (ZFC) formulé pour la première fois par Ernst Zermelo en 1904 [Zer04]. Il peut s'énoncer de la manière suivante :

Axiome 1. *Pour tout ensemble X d'ensembles non vides, il existe une fonction définie sur X, appelée fonction de choix, qui à chaque ensemble A appartenant à X associe un élément de cet ensemble A.*

$$\forall X \left[\emptyset \notin X \Longrightarrow \exists f : X \to \bigcup X \; \forall A \in X \; (f(A) \in A)\right]$$

Cet axiome du choix a été introduit dans le cadre de la démonstration du théorème du bon ordre, aussi appelé théorème de Zermelo qui affirme :

Théorème 7. *Tout ensemble peut être muni d'une structure de bon ordre, c'est à dire d'un ordre tel que toute partie non-vide admette un plus petit élément.*

Nous allons montrer qu'il y a équivalence entre l'axiome du choix et le théorème du bon ordre. C'est un résultat classique, redémontrons-le.

L'axiome du choix implique le théorème du bon ordre

Pour cette démonstration, nous emploierons le lemme de Zorn, équivalent à l'axiome du choix [Pey08]. Avant d'énoncer ce lemme, rappelons quelques définitions.

Une chaîne est définie comme suit :

Définition 8. *Soit X un ensemble muni d'un ordre $\leqslant$. On dit qu'une partie $C \subset X$ est une chaîne de X lorsque $\leqslant$ définit un ordre total sur C. On a donc $\forall x \neq y, x \in C, y \in C$, on a $x < y$ ou $y < x$.*

Un ensemble inductif se définit de la manière suivante :

Définition 9. *Un ensemble ordonné X est dit inductif lorsque chacune de ses chaînes admet au moins un majorant, c'est à dire un élément a tel que pour tout élément x de la chaîne, $x \leqslant a$.*

Enfin, le lemme de Zorn s'énonce comme suit :

> **Lemme 1.** *Tout ensemble inductif admet un élément maximal.*

Démonstration. Soit A un ensemble non-vide. Considérons les paires $(B, \leqslant_B)$, constituées des sous-ensembles $B \subset A$ et d'un bon ordre $\leqslant_B$ sur B.

On définit un ordre partiel sur l'ensemble X de toutes ces paires : $(B, \leqslant_B) \preceq (B', \leqslant_{B'})$ traduit que $B \subseteq B'$ et que $\leqslant_{B'}$ prolonge $\leqslant_B$. On entend par ce prolongement que $\leqslant_{B'}$ est le même bon ordre que $\leqslant_B$ mais valable sur tout l'ensemble B'.

Pour tout sous-ensemble bien ordonné $\{(B_i, \leqslant_{B_i})\}$ de X, c'est à dire pour toute chaîne de X, la paire $(\cup_i B_i, \cup_i \leqslant_{B_i})$ constitue un majorant. X est donc un ensemble inductif.

D'après le lemme de Zorn, X admet un élément maximal que nous notons $(M, \leqslant_M)$.

Supposons que $M \neq A$. Soit $a \in A \setminus M$. Nous pouvons définir un bon ordre sur l'ensemble $M \cup \{a\}$ en conservant le même ordre $\leqslant_M$ et en posant $\forall m \in M, m \leqslant a$. Nous avons prolongé le bon ordre et $M \subset M \cup \{a\}$ d'où $(M, \leqslant_M) \prec (M \cup \{a\}, \leqslant_{M \cup \{a\}})$. Donc $(M, \leqslant_M)$ n'est pas un élément maximal de X. Nous avons abouti à une contradiction.

Donc $M = A$ et $\leqslant_M = \leqslant_A$. A est bien muni d'un bon ordre.

Le théorème du bon ordre est vérifié.

□

Le théorème du bon ordre implique l'axiome du choix

Démonstration. Soit A un ensemble non-vide. Par le théorème du bon ordre, on peut munir cet ensemble d'un bon ordre.

Posons f la fonction qui associe à tout sous-ensemble S de A, son plus petit élément.

$$f(S) = \min(S)$$

Nous avons trouvé une fonction de choix sur A, ensemble quelconque. L'axiome du choix est donc vérifié.

□

Quel lien avec l'induction ?

Nous venons de montrer qu'il y a équivalence entre l'axiome du choix et le théorème du bon ordre. Cela signifie que si nous incluons dans notre système axiomatique l'axiome du choix, alors nous pouvons munir tout ensemble d'un bon ordre. Or, nous avons démontré plus tôt l'équivalence entre le principe de bon ordre et le principe d'induction. Ainsi, **l'axiome du choix permet en théorie d'appliquer le principe d'induction sur un ensemble quelconque.**

8 L'induction comme conséquence de l'axiome de l'infini

Résumé

Si dans l'arithmétique de Peano, le principe d'induction est un axiome, ce n'est pas le cas dans la théorie des ensembles (ZFC), il s'agit d'un théorème. Détaillons la manière d'aboutir à l'induction directement à partir des axiomes et en en utilisant le moins possible. Nous verrons que l'axiome de l'infini jouera un rôle primordial dans cette preuve.

La théorie de Zermelo-Fraenkel

La théorie des ensembles de Zermelo-Fraenkel (ZFC) dévelopée par les mathématiciens Ernst Zermelo et Abraham Fraenkel au début du XX$^{\text{ème}}$ siècle est le système axiomatique sur lequel se fonde l'ensemble des mathématiques.

Axiomes 2 (ZFC). *Lesl axiomes de ZFC s'énoncent comme suit [Kan] :*

1. *Axiome d'extensionabilité*

$$\forall A\, \forall B\ \ [\forall x\, (x \in A \Leftrightarrow x \in B) \Rightarrow A = B]$$

2. *Axiome de la paire*

$$\forall x\, \forall y\, \exists z\, (x \in z \vee y \in z)$$

3. *Axiome de la réunion*

$$\forall A\, \exists C\, \forall x\, [x \in C \Leftrightarrow \exists y\, (y \in A \wedge x \in y)]$$

4. *Axiome de l'ensemble des parties*

$$\forall A\, \exists C\, \forall x\, [x \in C \Leftrightarrow x \subset A]$$

5. *Axiome de l'infini*

$$\exists A\, [\emptyset \in A \wedge \forall x\, (x \in A \Rightarrow x \cup \{x\} \in A)]$$

6. *Schéma d'axiomes de compréhension*

$$\forall a_1 ... \forall a_p\, \forall A\, \exists B\, \forall x\, [x \in B \Leftrightarrow (x \in A \wedge P(x, a_1, ..., a_p))]$$

7. *Schéma d'axiomes de remplacement*

$$\forall a_1 ... \forall a_p\, \forall A\, [(\forall x, x \in A \Rightarrow (\exists! y\, P(x, y, A, a_1, ..., a_p))) \Rightarrow$$

$$\exists B\, (\forall x, x \in A \Rightarrow \exists y\, (y \in B \wedge P(x, y, A, a_1, ..., a_p)))]$$

8. *Axiome de fondation*

$$\forall A\ [A \neq \emptyset \Rightarrow \exists x\ (x \in A \wedge x \cap A = \emptyset)]$$

9. *Axiome du choix*

$$\forall X\ \left[\emptyset \notin X \Rightarrow \exists f : X \to \bigcup X\ \forall A \in X\ (f(A) \in A)\right]$$

Notre objectif n'est pas de dresser une démonstration du principe d'induction à partir de tous ces axiomes. En réalité, nous l'avons déjà fait de manière implicite en démontrant l'induction sur des ensembles bien ordonnés à partir du principe du bon ordre ainsi que l'induction sur des ensembles bien fondés grâce à l'axiome de fondation. En nous autorisant à employer l'axiome du choix, nous avons même démontré qu'il était possible de raisonner par induction sur tout ensemble en le munissant d'un bon ordre par le théorème du bon ordre.

Ici, notre but est de faire découler l'induction directement d'un nombre minimal d'axiomes. Pour ce faire, nous nous limiterons à la théorie initiale de Zermelo (Z) publiée en 1908 [Zer08], sans l'axiome du choix, c'est à dire les axiomes (1) à (6).

Construction des entiers naturels

Nous suivons la construction de Von Neumann [VN23] qui réprésente 0 par l'ensemble vide et et le successeur par la fonction $x \mapsto x \cup \{x\}$. L'ensemble $x \cup \{x\}$ peut être formé grâce à l'axiome de la réunion (3) et à l'axiome de la paire (2). Les entiers naturels sont donc les éléments x qui vérifient :

$$\forall A\ [\emptyset \in A \wedge \forall a\ (a \in A \Rightarrow a \cup \{a\} \in A)] \Rightarrow x \in A.$$

C'est l'axiome de l'infini (5) qui nous assure l'existence d'au moins un ensemble A. Et c'est le schéma d'axiomes de compréhension (6) qui nous permet de définir l'ensemble des entiers naturels comme l'ensemble des éléments qui vérifient la condition précédente. Aussi, l'axiome d'extensionabilité (1) affirme l'unicité de cet ensemble appelé $\mathbb{N}$.

L'induction en découle

La définition même de $\mathbb{N}$ fournit un principe d'induction sur les éléments de $\mathbb{N}$. En effet, soient P une propriété exprimée dans le langage de la théorie des ensembles et A l'ensemble des éléments qui vérifient P. Soit $S(x)$ le successeur de l'élément x. On peut écrire :

$$[P(0) \wedge \forall x \in \mathbb{N}\ (x \in A \Rightarrow S(x) \in A)] \Rightarrow \forall x \in \mathbb{N}, x \in A.$$

Nous avons abouti au principe d'induction à partir d'un nombre minimal d'axiomes, ceux de la thérie axiomatique de Zermelo tel qu'il l'a décrit en 1908 sans l'axiome du choix. Par ailleurs, il est utile de souligner que la propriété P est exprimée dans le langage de la théorie des ensembles plus fort que l'arithmétique de Peano.

9 Le problème des chapeaux à l'infini et son rapport à l'induction

Résumé

A partir de la connaissance d'un passé aussi proche que nous souhaitons, Hardin et Taylor proposent une stratégie nous assurant une prédiction exacte presque sûrement du présent, et même de quelques instants futurs. À travers différents problèmes de chapeaux, nous clarifions leur stratégie et retraçons les preuves de son efficacité. Enfin, nous faisons le lien avec le problème de l'induction qui nous a occupé au chapitre 1.

L'énoncé géneral du problème

Imaginons un directeur de prison qui place tous ses prisonniers en file indienne et fait porter à chacun d'eux un chapeau de couleur. Chaque prisonnier dans la file ne peut voir que les chapeaux des prisonniers devant lui. Il faut préciser que les prisonniers sont en nombre infini, qu'il y a une infinité de couleurs possibles et que le directeur est totalement libre de faire porter n'importe quelle couleur à n'importe quel prisonnier.

Si un prisonnier souhaite retrouver la liberté, il doit deviner la couleur de son chapeau. Il est évidemment totalement interdit aux prisonniers de communiquer dès lors que le directeur a posé les chapeaux mais ils peuvent tout de même s'accorder sur une stratégie au préalable [HT08a].

Combien de prisonniers pourront s'en sortir ? À priori, si un prisonnier s'en sortait, ce serait déjà un miracle. En effet, les chapeaux que les prisonniers voient n'ont aucun lien avec leur propre chapeau alors comment pourraient-ils deviner la couleur de leur chapeau à partir de ces observations qui semblent tout à fait inutiles ?

Hardin et Taylor ont trouvé une stratégie reposant sur l'axiome du choix qui permet :

- dans le cas où le nombre de prisonniers et de couleurs est dénombrable, de ne faire perdre qu'un nombre fini de prisonniers.
- dans le cas non-dénombrable, de ne faire perdre qu'un nombre infini dénombrable de prisonniers.

Dans les deux cas, le résultat va bien au-delà de nos espérances en faisant gagner la quasi-totalité des prisonniers. Clarifions donc cette stratégie et la preuve de son efficacité.

La stratégie Hardin-Taylor

Que ce soit dans cas dénombrable ou indénombrable, la stratégie proposée par Hardin et Taylor reste la même. Baptisée la μ-stratégie, elle est décrite comme suit :

> **Définition 10** (μ-stratégie). *[HT16] Soient* A *l'ensemble des joueurs,* K *l'ensemble des couleurs et* F *l'ensemble des fonctions qui associent à chaque joueur* $a \in A$ *une couleur* $k \in K$*. Soit* $[f]_a$ *la classe d'équivalence qui regroupe toutes les fonctions équivalentes à la fonction visible par l'agent* a, $f_{]a,+\infty[}$*. La notion d'équivalence diffère suivant que* A *et* K *sont dénombrables ou non, et sera donc précisée par la suite.*
>
> *Fixons un bon ordre* $\prec$ *sur* F *et pour chaque* $a \in A$*, écrivons* $\langle f \rangle_a$ *le plus petit élément de*

$[f]_a$. *La stratégie μ_a de l'agent a consiste alors à supposer que la véritable fonction $f \in F$ est la plus petite fonction selon l'ordre $\prec$ parmi celles qui peuvent s'accorder avec ce qu'il voit, et donner sa valeur à sa position. Plus précisément,* $\mu_a(f) = \langle f \rangle_a(a)$.

Il est utile de remarquer l'utilisation de l'axiome du choix dans cette stratégie. Il est indispensable pour pouvoir fixer un bon ordre sur un espace quelconque, ici sur l'ensemble des fonctions F, ce que nous avons démontré au chapitre 8.

Dans le cas dénombrable

Prenons le cas d'ensembles A (joueurs) et K (couleurs) comprenant une infinité dénombrable d'éléments. Les joueurs peuvent ainsi être numérotés à partir de 0. Définissons la notion d'équivalence dans ce cas.

Définition 11. *Deux fonctions f et g sont équivalentes si et seulement si l'ensemble* $\{x \mid f(x) \neq g(x)\}$ *des différences est fini.*

Ainsi, chaque joueur dans la file qui observe tous les chapeaux des joueurs devant lui peut déterminer dans quelle classe d'équivalence se trouve f puisque les seuls chapeaux qu'il ne peut pas observer sont en nombre fini. Par définition de la μ-stratégie, tous les joueurs s'accordant à prendre la même fonction dans cette classe d'équivalence, ils ne commettront qu'un nombre fini d'erreurs. En particulier, la μ-stratégie préconise de prendre le plus petit élément qui pourrait représenter la fonction "la plus simple" selon Hardin et Taylor [HT16].

Dans le cas indénombrable

Prenons maintenant le cas d'ensembles A et K comprenant une infinité indénombrable d'éléments. Il s'agit maintenant d'un continuum de joueurs. Pour la suite, nous parlerons moins de joueurs et de couleurs pour privilégier le cas général de fonctions sur les réels dans les réels ce qui nous permettra plus aisément de comprendre la stratégie comme un moyen de prédire le présent et le futur.

L'équivalence est ici définie comme suit :

Définition 12. *Soit $\lhd$ un ordre sur A. Deux fonctions f et g sont équivalentes selon la relation $\approx_t$ si et seulement si l'ensemble* $\{s \lhd t, s, t \in A \mid f(s) \neq g(s)\}$ *est au plus infini dénombrable.*

Ainsi, si les valeurs de la fonction f sont connues sur $]-\infty, t[$, la μ-stratégie nous permet de déterminer $f(t)$ de manière quasi-sûre.

En effet, posons $W_0 = \{t \in A \mid \langle f \rangle_t(t) \neq f(t)\}$ qui correspond à l'ensemble des erreurs commises en employant la μ-stratégie. Reprenons la démonstration de la dénombrabilité de cet ensemble [HT08b].

Démonstration. Montrons d'abord que W_0 est bien fondé en raisonnant par l'absurde.

Supposons qu'il existe une séquence infinie descendante telle que $\forall i \in \mathbb{N}$, $t_{i+1} \lhd t_i$ avec $\lhd$ transitive.

Par définition de l'équivalence, $\langle f \rangle_{t_i} \approx_{t_i} f$ d'où $\langle f \rangle_{t_i}(t_{i+1}) = f(t_{i+1})$.

De plus, comme $t_{i+1} \in W_0$, on a $\langle f \rangle_{t_{i+1}}(t_{i+1}) \neq f(t_{i+1})$.

Donc $\langle f \rangle_{t_i}(t_{i+1}) \neq \langle f \rangle_{t_{i+1}}(t_{i+1})$.

Ainsi, $\langle f \rangle_{t_i} \neq \langle f \rangle_{t_{i+1}}$.

Aussi, posons $\lhd t = \{s \in A \mid s \lhd t\}$. On a $\lhd t_{i+1} \subseteq \lhd t_i$. Les fonctions de $[f]_{t_i}$ doivent être équivalentes à f sur $]-\infty, t_i[$ soit sur un intervalle supplémentaire par rapport aux fonctions de $[f]_{t_{i+1}}$ qui correspond à $]t_{i+1}, t_i[$. D'où $[f]_{t_i} \subseteq [f]_{t_{i+1}}$.

On peut donc comparer les plus petits éléments de ces deux classes d'équivalence bien ordonnées : $\langle f \rangle_{t_{i+1}} \preceq \langle f \rangle_{t_i}$. Ayant démontré précédemment que $\langle f \rangle_{t_i} \neq \langle f \rangle_{t_{i+1}}$, on peut écrire l'inégalité stricte : $\langle f \rangle_{t_{i+1}} \prec \langle f \rangle_{t_i}$.

De cette manière, nous avons construit une chaîne infinie descendante pour l'ordre $\prec$ ce qui contradictoire puisqu'il s'agit d'un bon ordre. W_0 est donc bien fondé.

Si l'ordre $\lhd$ est choisi comme l'ordre $<$ avec $A = \mathbb{R}$, alors dans ce cas W_0 est un ensemble bien ordonné sur $\mathbb{R}$. Il y aura donc toujours un rationnel entre deux éléments de W_0, par densité de $\mathbb{Q}$ dans $\mathbb{R}$.

Donc W_0 est dénombrable.

□

Il est prouvé dans [HT08b] qu'en plus d'être dénombrable l'ensemble W_0 est nulle part dense. Il a donc, par exemple, encore moins de poids que les rationnels. La prédiction du présent en observant le passé est quasi-certaine dans ce sens.

La stratégie peut être renforcée en réduisant les données concernant le passé. Nul besoin de connaître tout le passé puisqu'il suffit de disposer d'un nombre infini non dénombrable de valeurs. Pour prédire $f(t)$, la connaissance de $[t - \epsilon, t[$ avec $\epsilon > 0$ aussi petit qu'on veut est suffisante.

Prédire le futur ?

Nous avons montré précédemment que la μ-stratégie était capable de prédire le présent. Mais pourrait-elle également prédire l'avenir ? Oui mais seulement sur un intervalle $[t, t+\epsilon[$ avec $\epsilon > 0$.

En effet, posons $W_1 = \{t \in A \mid \langle f \rangle_t \neq \langle f \rangle_{t'} \ \forall t \lhd t'\}$ qui représente l'ensemble des instants t où la valeur prise en adoptant la μ-stratégie diffère de toutes les valeurs que prendra par la suite la même stratégie. Il est utile de remarquer que $W_0 \subseteq W_1$. Et pour démontrer que l'ensemble W_1 est dénombrable, on procède exactement de la même manière que pour W_0.

Reprenons la démonstration de l'existence d'un ϵ tel que la stratégie prédise le futur jusque $[t, t+\epsilon[$ [HT08b].

Démonstration. Soit $t \in \mathbb{R} \setminus W_1$. Par définition de W_1, il existe $t' > t$ tel que $\langle f \rangle_t = \langle f \rangle_{t'}$.

On a donc : $\forall u < t', \langle f \rangle_t(u) = \langle f \rangle_{t'}(u) = f(u)$, par définition de l'équivalence.

Ainsi, $\langle f \rangle_t$ et f coïncident sur $]-\infty, t'[$ et donc sur $[t, t'[= [t, t+\epsilon[$. La μ-stratégie permet une bonne prédiction sur $\mathbb{R} \setminus W_1$ avec W_1 un ensemble dénombrable. □

De quelle sorte d'induction s'agit-il ?

Depuis le 2[ème] chapitre, nous avons traité de l'induction mathématique sous des formes variées. Toutes les formes d'induction présentées ont en commun la structure de leurs hypothèses et la forme de leur conclusion. Il s'agissait toujours, en partant d'une propriété héréditaire et vraie pour au moins un élément, de montrer que cette propriété était vraie pour tous les éléments de l'ensemble. Il ne peut pas s'agir ici d'induction mathématique comme nous l'avons entendue jusqu'à présent. Il s'agit d'une vériable prédiction mathématique et pas seulement d'un outil de preuve.

En revanche, il est possible de rapprocher ces prédictions de l'induction scientifique décrite au 1[er] chapitre. A partir de cas particuliers observés de l'univers qui n'ont pas nécessairement de lien entre eux, nous prédisons par la μ-stratégie le résultat de cas non-observés de ce même univers.

Le problème de l'induction est-il résolu ?

Le problème de l'induction posé par Hume pourrait-il être résolu grâce à la μ-stratégie ? C'est la question que ce sont notamment posé Alexander George et Alexander Paseau. Exposons leurs principaux arguments.

Alexander George

L'utilisation de l'axiome du choix nous empêche de pouvoir employer la μ-stratégie. La fonction de choix existe mais nous ne pouvons pas la construire. Il nous est impossible d'ordonner des collections de fonctions et donc de réaliser une prédiction juste. Cette non-constructivité de la solution d'Hardin-Taylor nous oblige à distinguer dans le problème de l'induction deux sous-problèmes :

- Peut-on justifier l'existence d'une stratégie qui ne se fonde que sur des informations du passé pour prédire l'avenir ?
- Peut-on justifier nos prédictions qui utilisent cette stratégie ?

L'existence d'une telle stratégie est bien démontrée mathématiquement mais quelle justification apporter à nos prédictions qui ne se fondent que sur des observations sans aucun lien ni entre elles ni avec les futures observations ?

Enfin, George met en lumière deux affirmations vraies qui semblent paradoxales de prime abord :

1. Toute fonction f est telle que, pour presque tous les $t \in \mathbb{R}$, la μ-stratégie prédit correctement la valeur de $f(t)$.
2. Tout point $t \in \mathbb{R}$ est tel que, pour presque toutes les fonctions f, la μ-stratégie prédit incorrectement la valeur de $f(t)$.

La 1ère affirmation a été démontrée par Hardin et Taylor et la 2nde traduit simplement le fait que la probabilité de prédire correctement un nombre réel choisi aléatoirement est nulle. Elles ne se contredisent pas puisque les quantificateurs sont inversés mais cela reste pour le moins surprenant. En fait, cela montre que la μ-stratégie n'est pas efficace localement, c'est à dire à un certain t seulement. Elle ne prend son sens que dans la globalité d'une infinité d'instants. Le succès des prédictions de la μ-stratégie apparait encore plus mystérieuse [Geo07].

Alexander Paseau

Comme George, Paseau insiste sur la non-constructivité de la μ-stratégie et sur son impossible application : une fonction de choix inaccessible, le traitement d'une infinité non-dénombrable de valeurs, la connaissance parfaite de l'univers et encore les états de l'univers encodés chacun par un unique nombre réel.

Paseau fait d'ailleurs la distinction entre plusieurs degrés de non-constructivité. Pour illustrer le 1er degré, il prend l'exemple de la simple démonstration de l'existence d'un nombre infini indénombrable de nombres transcendants (réels non-algébriques, c'est à dire qui ne sont pas solutions d'une équation polynomiale à coefficients rationnels). Cette existence est déduite du fait qu'il n'y a qu'un nombre infini dénombrable de nombres algébriques. Même si l'on peut construire des nombres transcendants (comme e), la preuve précédente ne nous en donne pas les moyens.

Pour le 2ème degré, il donne l'exemple de la démonstration du fait qu'un irrationnel mis à la puissance d'un irrationnel peut être rationnel. Soit $\sqrt{2}^{\sqrt{2}}$ est rationnel et dans ce cas la paire $(\sqrt{2}, \sqrt{2})$ est un exemple de ce que l'on veut prouver ou alors il est irrationnel et c'est la paire $(\sqrt{2}^{\sqrt{2}}, \sqrt{2})$ qui fonctionne. La preuve nous donne plusieurs possibilités mais nous ne savons pas laquelle est la bonne.

Un dernier degré de preuve non-constructive concerne celles dont les instances ne peuvent en aucun cas être exhibées. En faisant appel à un bon ordre sur les réels, la preuve de Hardin et Taylor entre dans cette dernière catégorie. En effet, il est démontré qu'un bon ordre sur les réels ne peut pas être défini dans le langage de ZFC. De ce fait, la μ-stratégie ne peut même pas être définie dans ce langage. C'est pour cette raison que Paseau scinde la première question posée par George en deux :

— Peut-on justifier **l'existence d'une fonction** qui ne se fonde que sur des informations du passé pour prédire l'avenir ?
— Peut-on justifier **l'existence d'une stratégie** qui ne se fonde que sur des informations du passé pour prédire l'avenir ?

Il fait donc la disctinction entre fonction et stratégie, et selon lui, la solution d'Hardin et Taylor ne répond positivement que dans le cas d'une fonction [Pas08].

Alexander Paseau appuie également sur l'importance de la structure temporelle qui n'est aujourd'hui pas une question tranchée en physique. Si le temps est discret, alors il est nécessaire de remonter à tous les états de l'univers dans le passé, de $-\infty$ à t. Or, les modèles actuels de la physique affirment l'existence d'un instant 0 ce qui rend impossible l'application de la μ-stratégie. Si le temps est continu, alors c'est l'uniformité de la nature, mentionnée au chapitre 1, qui peut toujours poser problème. La structure du temps pourrait changer à tout moment.

Enfin, il remet en question l'utilité de prédire le présent alors qu'on pourrait simplement l'observer et même dans le cas où la μ-stratégie nous laisse entrevoir un peu du futur, il est impossible de savoir jusque quand la prédiction est valide [Pas11]. Si la poule de Russell appliquait la stratégie et qu'elle obtenait un résultat positif ("personne ne lui tordra le cou") pour l'heure qui suit, serait-elle pour autant sereine ? Non, puisqu'elle ne sait pas où s'arrête la validité de la prédiction : 1 heure, 1 minute, 1 seconde... ?

10 Réflexions, problèmes ouverts et perspectives

Déterminer la structure du temps : de l'induction...

Comme pour tous les autres résultats établis en physique, si on veut montrer que le temps est continu, il faudra le faire par un raisonnement inductif. En ayant montré que le temps est continu... on pourra possiblement affirmer que la stratégie de prédiciton de Hardin-Taylor est possible, l'induction serait alors justifiée. Mais notre raisonnement est circulaire : l'induction ne peut pas justifier l'induction !

Tentative infructueuse de prédire la prédiction

Dans le cas continu, on observe les valeurs d'une fonction f jusque t'. On utilise la stratégie Hardin-Taylor pour prédire les valeurs de f sur $[t, t+\epsilon[$ avec $t+\epsilon < t'$. Nous appelons g la fonction prédite. Grâce à notre observation de f jusque t' (aussi grand que l'on veut), on pourra déterminer l'instant $t+\epsilon$ qui correspond à l'instant où la prédiction devient inexacte. Posons h une fonction définie comme suit :

— Si $f(x) = g(x)$, alors $h(x) = 1$
— Sinon, $h(x) = 0$

Nous connaissons la fonction h sur $[t, t']$. Elle est par ailleurs constante égale à 1 sur $[t, t+\epsilon[$ Essayons de prédire la fonction h, c'est à dire de prédire la validité de la prédiction de f. Prenons un $\epsilon_2 < \epsilon$. Grâce à Hardin-Taylor, nous pouvons prédire h sur $[t+\epsilon_2, t+\epsilon_2+\epsilon_3[$.

Connaissant la fonction h, nous pouvons comparer les valeurs à la prédiction et ainsi déterminer ϵ_3 de la même manière que précédemment. Si $\epsilon_2 + \epsilon_3 < \epsilon$, h doit être constante égale à 1 (sauf sur une infinité dénombrable).

Rien de pertinent...Dommage...

Une induction à la Goodman ?

Même si les données que l'on a correspondent à une infinité de chapeaux de couleur verte, notre stratégie ne s'en préoccuppe pas vraiment pour la suite et pourra prédire n'importe quelle couleur. On peut rapporcher cela de la nouvelle énigme de Goodman que l'on a évoqué au chapitre 1. Si jusqu'à présent, toutes les émeraudes observées sont vertes, doit-on induire que les futurs émeraudes seront vertes également ou qu'elles deviendront bleu à un certain moment ?

Poincaré répondrait qu'il faut aller au plus simple et opter pour la première option, c'est le rasoir d'Occam. D'ailleurs, lorsque Hardin et Taylor munissent l'ensemble des fonctions possibles d'un bon ordre, il font référence à ce principe : le plus petit élément serait l'élément le plus simple. Mais ce n'est qu'une volonté, l'axiome du choix ne nous permet pas de définir le bon ordre que l'on préfère, mais nous assure seulement l'existence d'un bon ordre. Avec la μ-stratégie, nous ne sommes donc pas assurés d'aller au plus simple.

Indépendance des évènements

Peut-on vraiment parler d'indépendance des évènements alors que c'est "grâce" à l'observation d'une infinité d'autres évènements que la stratégie peut être efficace ? Il est important de le savoir si l'on souhaite parler de probabilités.

Utile uniquement pour les fonctions discontinues

Dans le cas indénombrable, pour les fonctions continues, prédire le présent en employant la μ-stratégie ne sert pas à grand chose. Il suffit d'utiliser la notion de limite. La stratégie n'a d'intérêt que dans le cas de fonctions discontinues.

Du non-standard chez Hardin-Taylor ?

Prouvons par induction la propriété $P(n)$: "il existe une configuration gagnante de chapeaux (seulement un nombre fini d'erreurs) telle que les n premières prédictions soient fausses".

— Il existe une configuration de chapeaux telle que la $1^{\text{ère}}$ prédiction soit fausse puisque la stratégie autorise un nombre fini d'erreurs, ici une seule ($P(1)$ est vrai).
— Supposons qu'il existe une configuration de chapeaux telle que les n premières prédictions soient fausses. Alors il existe une configuration de chapeaux telle que les $n+1$ premières prédictions soient fausses, puisque la stratégie autorise un nombre fini d'erreurs, ici $n+1$.

Par induction, on conclut que pour tout n, les n premières prédictions peuvent être fausses...

Cela nous rappelle la preuve par récurrence incorrecte que $\forall n, n$ est fini du chapitre 5. Le fait qu'il n'y ait qu'un nombre fini d'erreurs signifie qu'il existe un n^* tel que toutes les prédicitons au-delà de n^* sont justes. Mais où est ce n^* ? Se pourrait-il qu'il s'agisse d'un nombre non-standard ?

Si traiter l'infini est possible alors le problème de l'induction est moindre ?

Avec Hardin-Taylor, il faut traiter l'infini et même faire une infinité de choix dans des ensembles infinis. Or, autoriser cet usage de l'infini élimine déjà une partie du problème de l'induction puisqu'il nous laisse la possibilité de tester une infinité de cas pour vérifier ou réfuter une loi physique. Encore faudrait-il pouvoir dresser la liste de tous les cas envisageable, ce qui n'est pas toujours possible.

Un nouveau paradoxe du menteur dans la file infinie de joueurs

Imaginons que chaque agent donne sa prédiction au chef des chapeaux. Après avoir fait sa prédiction, chaque agent peut maintenant regarder la véritable couleur de son chapeau. Tout le monde, simultanément, dit "vrai" si sa prédiction était vraie, et "faux" sinon. Sachant que chaque agent voit tous les chapeaux devant lui, il peut savoir si l'un des agents devant a menti.

Imaginons que tous les agents disent : "tous mentent devant moi". Raisonnons par l'absurde. Supposons qu'un agent n dise la vérité. Alors l'agent $n + 1$ ment puisque $n + 1 > n$. Il y a donc un agent $m > n + 1$ qui dit la vérité. Or $m > n$, donc l'agent m ment. Nous avons abouti à une contradiction : l'agent m ne peut pas à la fois dire la vérité et mentir. Donc tous les agents mentent.

Mais si tous les agents mentent alors en disant "tous mentent devant moi", ils disent tous la vérité. C'est un paradoxe, publié par Yablo en 1993.

Il n'y a pas ici d'auto-référence et pourtant nous avons abouti à un paradoxe ! Une infinité d'affirmations portant sur la véracité d'une infinité d'affirmations, c'est ce qui semble être la cause du paradoxe.

La μ-stratégie ne ressemble pas à l'induction scientifique

L'induction telle qu'elle est pratiquée en sciences ne ressemble que de loin à l'approche de la μ-stratégie. Jamais un physicien ne tente de se transformer en démon de Laplace qui connait tout le passé de l'univers et prédit l'avenir de l'Univers. Ce qu'il fait, par exemple pour la loi de la gravité, c'est observer un nombre fini de cas où la loi s'applique sous différentes conditions (sans jamais observer un cas où elle ne s'applique pas) et d'induire que la loi est vraie pour tous les cas, et donc même pour l'infinité de cas qui n'a pas été observée. Le scientifique veut en apprendre davantage sur la fonction f qui régit le monde et pas seulement connaître des valeurs non justifiées prises par f.

Des applications aux divers problèmes de chapeaux ?

Quelles pourraient-être les applications des solutions aux divers problèmes de chapeaux (dans le cas où elles sont constructives) ? On peut penser à la transmission de données. Il y a un toujours un risque de distorsion et si un 0 se transforme en 1 accidentellement, il faut repérer l'erreur. On cherche alors à ajouter un nombre minimal de bits de correction et certain problèmes de chapeaux (finis) peuvent mener à des solutions. Certains y trouvent également une utilité dans l'approximation de fonctions binaires par des fonctions polynomiales ou encore lorsqu'il s'agit de réduire des séquences à priori aléatoires [Krz10].

Difficultés des probabilités à l'infini

Imaginons une loterie infinie avec autant de tickets que d'entiers naturels. On se demande quelle est la probabilité que le ticket gagnant porte un numéro pair. On est tenté de répondre 50% et pourtant, n'y a-t-il pas autant de nombres pairs que de nombres à l'infini ?

On ne peut pas se calquer sur notre intuition du fini pour parler d'infini. Est-il même possible d'attribuer une probabilité à un ticket tout en conservant une loterie qui soit juste ? C'est la question posée par Sylvia Wenmackers et Leon Horsten. Pour ce faire, ils font appel à des ensembles non-standards ce qui leur permet d'assigner une probabilité infinitésimale à chaque ticket. [WH13]

De la même façon, pour imaginer des prédictions probabilistes dans le problème des chapeaux, ne pourrait-on pas adopter, notamment dans le cas dénombrable, une approche similaire ? Notons que ces probabilités infinitésimales qui correspondent à des hyperrationnels ne peuvent exister sans l'axiome du choix...

Pas de problème de l'induction scientifique dans les mathématiques

A l'origine, la dernière phrase du chapitre 1 concernant le problème de l'induction, affirmait que ce problème ne se posait pas en mathématiques. Pourquoi n'y aurait-il pas de tels problèmes en mathématiques ?

L'induction scientifique n'est utilisée par les mathématiciens qu'avant de tenter de prouver une affirmation. Si après de nombreuses tentatives aucun contre-exemple n'est trouvé, alors on peut plus facilement croire en cette affirmation et se lancer dans sa démonstration. Ce raisonnement inductif n'est utile que pour formuler des conjectures et non des théorèmes, pour fonder des croyances et non pas pour établir des vérités. Dans son article sur le problème de l'induction dans les mathématiques [Bak07], Alan Baker va même plus loin en affirmant que même cette croyance n'est pas toujours fondée sur une énumération de cas valides.

Reprenons l'exemple de la conjecture de Goldbach. Elle a été vérifiée jusque 4.10^{18} et la quasi-totalité des mathématiciens s'accordent à dire qu'elle doit être vraie. On pourrait donc penser que cette croyance se fonde sur les tests qui ont été menés. Une autre conjecture qui affirme que tous les nombres parfaits sont pairs (nombre égal à la somme de ses diviseurs sans compter lui-même) ne fait quant à elle pas consensus. Pourtant, elle est équivalente à affirmer que tous les nombres impairs sont imparfaits et il a été prouvé qu'un nombre impair imparfait devait nécessairement être supérieur à 10^{300}. L'énumération est 10^{282} plus importante que dans le cas de la conjecture de Goldbach tandis que la croyance en sa véracité est moindre.

Le raisonnement inductif en sciences n'est certainement pas une manière d'établir des théorèmes en mathématiques et il est même légitime de douter de son utilité pour fonder la croyance des mathématiciens.

L'importance de la non-mesurabilité pour prédire

Prenons une variante du problème des chapeaux, d'ailleurs mentionnée dans le livre d'Hardin et Taylor. Les joueurs sont en nombre infini dénombrable, il y a uniquement 2 couleurs de chapeaux possibles et chaque joueur entend la prédiction des joueurs derrière lui. Jalex Stark prouve que si $(f_i)_{i\in\mathbb{N}}$ correspondant à la μ-stratégie appliquée par le joueur i, est une stratégie gagnante pour ce problème des chapeaux, alors f_1 est une fonction non-mesurable [Sta16]. f_i correspond

La démonstration se fait par par l'absurde en supposant que la fonction f_1 est mesurable, c'est à dire que les probabilités que $f_1 = 0$ et $f_1 = 1$ sont bien définies et par ailleurs égales par symétrie. La contradiction vient de l'application de la loi du zéro-un de Kolmogorov affirmant que tout événement dont la réalisation dépend d'une suite de variables aléatoires indépendantes mais ne dépend d'aucun sous-ensemble fini de ces variables est soit presque sûrement réalisé, soit presque sûrement non réalisé, c'est-à-dire que sa probabilité est de 0 ou 1. C'est effectivement le cas dans le problème des chapeaux dénombrable ce qui signifie que $P(f_1 = 0) = P(f_1 = 1) = 0$ ou 1. On aboutit à une contradiction. La stratégie du premier joueur n'est pas mesurable.

Peut-être pourrait-on généraliser cette nécessité d'employer des fonctions non-mesurables pour que le problème des chapeaux traditionnel ait une solution ? Ou si on se place dans une théorie où tous les ensembles sont mesurables, ne peut-on pas montrer qu'aucune stratégie n'est gagnante ? D'ailleurs, le théorème de Banach-Tarski, tout aussi surprenant, requiert l'existence de fonctions non-mesurables. Les morceaux de la boule initiale nécessaires à la construction de deux boules identiques à celle-ci ne sont pas mesurables.

La μ-stratégie ou la construction d'un ensemble de Vitali

Dans la μ-stratégie d'Hardin-Taylor, l'ensemble des choix des joueurs, c'est à dire l'ensemble composé des fonctions représentatives de chaque classe d'équivalence est un ensemble non-mesurable. Pour le montrer, il suffit de remarquer que sa construction suit exactement celle de l'ensemble de Vitali lui-même non-mesurable. Chez Vitali, deux nombres réels appartiennent à la même classe d'équivalence si la différence entre les deux est un nombre rationnel. L'ensemble qui porte son nom regroupe les représentants de chaque classe d'équivalence. Chez Hardin-Taylor, dans le cas dénombrable par exemple, deux fonctions appartiennent à la même classe d'équivalence si elles ne diffèrent qu'un nombre fini de fois. On ne peut donc pas attribuer de probabilité aux évènements "il y aura n mauvaises predictions" ou "la $n^{\text{ème}}$ prédiction est fausse".

Plus choquant que Banach-Tarski ?

Il est évident qu'en réalité, on ne peut pas construire deux boules identiques à partir d'une seule, le volume ne serait pas conservé. C'est peut-être pour ça que Banach-Tarski nous pose moins de problèmes qu'Hardin-Taylor : nous savons que les découpages de la boule ont lieu dans un monde mathématique, la démonstration faisant appel à l'axiome du choix, elle est non-constructive. Il en est de même pour Hardin-Taylor et pourtant on a davantage envie de faire le parallèle avec la réalité tant il est tentant de vouloir prédire l'avenir...

A Équivalence entre l'induction simple et forte

Résumé

Pour montrer l'équivalence entre l'induction et le principe du bon ordre, il est nécessaire de montrer l'équivalence entre deux formes d'induction : l'induction simple et forte. En voici la démonstration.

Les énoncés des principes

Principe 2 (Induction simple). *Si pour une propriété P concernant les entiers naturels n, la suite est vraie :*

1. *la propriété est vraie pour $n = 0$,*
2. *si la propriété est vraie pour $n = k$ alors elle est aussi vraie pour $n = k + 1$, $(*)$*

alors la propriété P est vraie $\forall n \in \mathbb{N}$.

Principe 3 (Induction forte). *Si pour une propriété P concernant les entiers naturels n, la suite est vraie :*

1. *la propriété est vraie pour $n = 0$,*
2. *si la propriété est vraie $\forall n \leqslant k$ alors elle est vraie pour $n = k + 1$, $(**)$*

alors la propriété P est vraie $\forall n \in \mathbb{N}$.

L'induction forte implique l'induction simple

Démonstration. Supposons que le principe d'induction forte est vérifié.

"$P(n)$ est vrai $\forall n \leqslant k$" implique "$P(k)$ est vrai" puisque $k \leqslant k$. Ainsi, si la proposition $(*)$ est vraie alors la proposition $(**)$ l'est également.

Donc si les propositions (1) et (2) du principe d'induction simple sont vraies, alors les propositions (1) et (2) du principe d'induction forte le sont aussi sachant que la proposition (1) est identique pour les deux. Pour toute propriété P qui vérifie les propositions (1) et (2) du principe d'induction simple, nous pouvons conclure par induction forte que P est vraie $\forall n \in \mathbb{N}$.

Il s'agit bien là de la conclusion du principe d'induction simple. Le principe d'induction simple est donc aussi vérifié. □

L'induction simple implique l'induction forte

Démonstration. Supposons que le principe d'induction simple est vérifié.

Soit une propriété P telle que les propositions (1) et (2) de l'induction forte sont vraies. Soit $Q(k)$, "$P(n)$ est vrai $\forall n \leqslant k$".

Nous voulons montrer que $P(n)$ est vrai $\forall n$, c'est à dire que $Q(n)$ est vrai pour $\forall n \in \mathbb{N}$.

Comme $P(0)$ est vrai, $Q(0)$ l'est aussi.

De plus, nous savons que si $P(n)$ est vrai $\forall n \leqslant k$ alors $P(k+1)$ est vrai. En d'autres termes, si $Q(k)$ est vrai alors $P(k+1)$ est vrai. Or, la combinaison de la validité de $Q(k)$ et de $P(k+1)$ implique la validité de $Q(k+1)$. Donc si $Q(k)$ est vrai alors $Q(k+1)$ l'est également.

Nous pouvons conclure par induction simple que $Q(n)$ est vrai pour $\forall n \in \mathbb{N}$. Donc $P(n)$ est vrai $\forall n \in \mathbb{N}$.

Il s'agit bien là de la conclusion du principe d'induction forte. Le principe d'induction forte est donc aussi vérifié.

□

Il y a donc bien équivalence entre le principe d'induction simple et forte.

Références

Ana02. Bhupinder Singh Anand. Beyond Gödel. 2002.

Bak07. Alan Baker. Is there a problem of induction for mathematics ? 2007.

Bon04. Jacqueline Boniface. Poincaré et le principe d'induction. *Philosophiques*, 31(1) :134–149, 2004.

Cag15. Pierre Cagne. Le nec plus ultra du théorème de compacité. 2015.

Col09. Pierre Colmez. *Eléments d'analyse et d'algèbre (et de théorie des nombres)*. Editions Ecole Polytechnique, 2009.

Coz09. M. Cozic. Le raisonnement scientifique, 2. Les problèmes de l'induction et de la confirmation (i), 2009.

Dav09. ISAAC Davis. An introduction to nonstandard analysis. 2009.

Ded65. Richard Dedekind. Was sind und was sollen die zahlen ? In *Was sind und was sollen die Zahlen ? Stetigkeit und Irrationale Zahlen*, pages 1–47. Springer, 1965.

Del02. André Deledicq. L'analyse non-standard, théorie des ordres de grandeur. *Tangente Hors série*, (13), 2002.

Del14. Jean-Paul Delahaye. Indécidables utiles et inutiles. *Pour la Science*, (442), Août 2014.

eSHP14. Tomàs Oliveira e Silva, Siegfried Herzog, and Silvio Pardi. Empirical verification of the Even Goldbach Conjecture and Computation of Prime Gaps up to 4.10^{18}. *Math. Comput.*, 83(288) :2033–2060, 2014.

Geo07. Alexander George. A proof of induction ? 2007.

Göd29. 1906-1978 [VerfasserIn] Gödel, Kurt. *Über die Vollständigkeit des Logikkalküls*. 1929.

Göd92. Kurt Gödel. *On formally undecidable propositions of Principia Mathematica and related systems*. Courier Corporation, 1992.

Hen18. Leah Henderson. The Problem of Induction. *The Stanford Encyclopedia of Philosophy*, 2018.

HT08a. Christopher S Hardin and Alan D Taylor. An introduction to infinite hat problems. *The Mathematical Intelligencer*, 30(4) :20–25, 2008.

HT08b. Christopher S Hardin and Alan D Taylor. A peculiar connection between the axiom of choice and predicting the future. *The American Mathematical Monthly*, 115(2) :91–96, 2008.

HT16. Christopher S Hardin and Alan D Taylor. *Mathematics of Coordinated Inference*. Springer, 2016.

Kac09. Tarik Kaced. Tri : De la semi-décidabilité incalculable à l'indépendance. 2009.

Kan. Akihiro Kanamori. ZFC - Encyclopedia of Mathematics. URL : `http://www.encyclopediaofmath.org/index.php?title=ZFC&oldid=19298`.

Kem58. John G. Kemeny. Undecidable problems of elementary number theory. *Mathematische Annalen*, 135, 1958.

Krz10. Marcin Krzywkowski. On the hat problem, its variations, and their applications. *Annales Universitatis Paedagogicae Cracoviensis Studia Mathematica*, 9(1) :55–67, 2010.

Lad69. Jean Ladrière. Le théorème de Löwenheim-Skolem. *Cahiers pour l'analyse*, 10 :108–130, 1969.

Liv. Equivalence, order, and inductive proof - url : `http://web.vu.lt/mif/e.manstavicius/wp-content/uploads/2017/09/Hein-DM-4.pdf`.

LY99. Igor LY. Goodman, Poincaré et la nouvelle énigme de l'induction. *Philosophiques*, 1999.

MP89. Michael Makkai and Robert Paré. *Accessible categories : the foundations of categorical model theory*, volume 104. American Mathematical Soc., 1989.

Mye13. Andrew Myers. Well-founded induction, 2013.

Pas08. Alexander Paseau. Justifying induction mathematically : strategies and functions. *Logique et Analyse*, 51(203) :263–269, 2008.

Pas11. Alexander Paseau. Proving induction. *Australasian Journal of Logic*, 10 :1–17, 2011.

Pey08. Rémy Peyre. Le lemme de Zorn. 2008.

Poi02. Henri Poincaré. *La science et l'hypothèse*. 1902.

Rag16. Giuseppe Raguní. A common misconception about the categorical arithmetic. *arXiv preprint arXiv :1602.03389*, 2016.

Ram95. Olivier Ramaré. On šnirel'man's constant. *Annali della Scuola Normale Superiore di Pisa-Classe di Scienze*, 22(4) :645–706, 1995.

Sal88. Jean Michel Salanskis. L'analyse non standard et la tradition de l'infini. *Revue d'histoire des sciences*, 41(2) :157–207, 1988.

Sko34. Th Skolem. Über die nicht-charakterisierbarkeit der zahlenreihe mittels endlich oder abzählbar unendlich vieler aussagen mit ausschliesslich zahlenvariablen. *Fundamenta mathematicae*, 23(1) :150–161, 1934.

Sta16. Jalex Stark. A countable hat game and the importance of measurability. 2016.

VN23. John Von Neumann. Zur Einführung der transfiniten Zahlen. *Acta Litterarum ac Scientiarum Regiae Universitatis Hungaricae Francisco-Josephinae, sectio scientiarum mathematicarum*, 1 :199–208, 1923.

WH13. Sylvia Wenmackers and Leon Horsten. Fair infinite lotteries. *Synthese*, 190(1) :37–61, 2013.

Zer04. E. Zermelo. Beweis, daÃ§ jede menge wohlgeordnet werden kann. (aus einem an herrn hilbert gerichteten briefe). *Mathematische Annalen*, 59 :514–516, 1904.

Zer08. E. Zermelo. Untersuchungen über die Grundlagen der Mengenlehre. I. *Mathematische Annalen*, 65 :261–281, 1908.